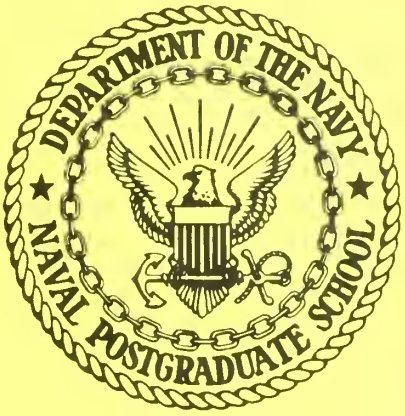
NAVAL POSTGRADUATE SCHOOL

Monterey, California



THE INFLUENCE OF DATA ORDER ON THE
LIEBERMAN-ROSS METHOD

by

Arthur L. Schoenstadt

April 1976

Prepared for:
Chief of Naval Research, Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral Isham Linder
Superintendent

Jack R. Borsting
Provost

The work reported herein was supported by the Foundation Research Program of the Naval Postgraduate School with funds provided by the Chief of Naval Research.

This report was prepared by:

Robert R. Fossum
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS-53Zh76047	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* THE INFLUENCE OF DATA ORDER ON THE LIEBERMAN-ROSS METHOD		5. TYPE OF REPORT & PERIOD COVERED Final-October 1975-December 1975 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Arthur L. Schoenstadt		8. CONTRACT OR GRANT NUMBER(s) N6685676WR00012
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Chief of Naval Research Arlington, VA 22217		12. REPORT DATE April 1976 13. NUMBER OF PAGES 42
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Lieberman-Ross Method
Reliability
Simulation

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

A statistically exact procedure for producing lower bounds on the reliability of a parallel system of independent, exponentially failing components has been produced by Lieberman and Ross. This method uses individual component failure data to estimate system characteristics. The method has the drawback that permuting the order in which components fail (but not the inter-failure times), can alter the estimated bound, i.e., the estimates are Data Order Dependent. In this paper we investigate this dependence of the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

20. (cont'd)

estimates, and consider the effects of a prior ordering.

TABLE OF CONTENTS

I. INTRODUCTION

The reliability of a system composed of N statistically independent components, is defined:

$$R(T_o) = \text{Probability } (T \geq T_o)$$

where T denotes the time of first system failure. A significant practical problem is that it is often reasonably easy, and comparatively inexpensive, to determine information about the reliability of separate components, denoted $R_i(T_o)$, $i=1,2, \ldots, N$, but fairly difficult and extremely expensive to determine total system reliability directly, since failure testing often destroys the system.

In the most common analytically treated case each of the component types has an assumed exponential failure rate, i.e.,

$$R_i(T_o) = e^{-\lambda_i T_o}, \quad i=1,2, \ldots, N, \tag{1}$$

the system reliability is given by

$$R(T_o) = e^{-\lambda T_o}, \quad \lambda = \sum_{i=1}^{N} \lambda_i \tag{2}$$

and a statistically exact procedure for obtaining an upper bound on λ from data on component failures, called the LR procedure, has been developed by Lieberman and Ross [1]. In this procedure k_i items of each type component, with individual observed failure times T_{ij}, $i=1,2, \ldots, N$ and $j=1,2, \ldots, k_i$ are tested. (We shall assume without loss of generality that $k_j \leq k_{j+1}$, $j=1,2, \ldots, N-1$.) A time U is defined:

$$U = \min_i \left\{ \sum_{j=1}^{k_i} T_{ij} \right\} . \tag{3}$$

i.e., U is the cumulative time at which one first exhausts all the components of one type. The number of each type of component which has been used up to this time is given by:

$$n_i = \{\text{largest } j \leq k_i \mid \sum_{\ell=1}^{j} T_{i\ell} \leq U_{min}\}. \tag{4}$$

Then, if

$$K = \sum_{i=1}^{N} n_i \tag{5}$$

(i.e., total components used), Lieberman and Ross showed that $2\lambda U$ follows the χ^2_{2K} distribution; hence upper bounds for λ can be estimated.

A major drawback of the LR technique is that in using K as defined by (4)-(5), one discards the "information" known about the

$$NL = \sum_{i=1}^{N} \{k_i - n_i\} \tag{6}$$

components which do not fail by the time U.

An immediate consequence of (4) is that the LR procedure is Data-Order Dependent. By this we mean that permuting the second subscript on T_{ij} (i.e., in essence permuting the order in which the failures are observed) can alter n_i, and hence the estimated bound. For example, consider the following hypothetical observed inter-failure times in two tests:

$$\text{Test A} \begin{cases} \text{Component \#1:} & T_{11} = 0.75 \ , \ T_{12} = 0.25 \\ \text{Component \#2:} & T_{21} = 0.80 \ , \ T_{22} = 1.25 \end{cases}$$

$$\text{Test B} \begin{cases} \text{Component \#1:} & T_{11} = 0.25 \ , \ T_{12} = 0.75 \\ \text{Component \#2:} & T_{21} = 1.25 \ , \ T_{22} = 0.80 \end{cases}$$

Observe the data are identical except for their order (i.e., which values are associated with which second subscripts), and $U = 1.00$ in both tests, however $K = 3$ in Test A but only $K = 2$ in Test B. This leads to upper bounds (at the 95% confidence level) for λ of 6.30 and 4.75 respectively. This dependence on the order in which the second subscript of the T_{ij} are assigned causes a potentially wide variance in the bounds for λ for the identical set of individual component failure data. This can lead to significant practical problems, for often interfailure data on individual components is presented with no reference as to the order in which the failures occurred. Thus there is no preferred ordering, and the analyst is faced with the formidable task of deciding on the "best" way to assign the T_{ij} .

We shall investigate, using analytic and simulation techniques, the impact of this ordering, and other parameters, on the mean and variance of the bounds produced for λ .

II. ANALYSIS OF THE LR METHOD - TWO COMPONENT CASE

Each application of the LR procedure to a given set of data involves a simultaneous observation of two random variables - the discrete random variable K, and the continuous random variable U. Given the confidence level, α, desired, then the estimated bound for λ,

$$\lambda_b = \frac{\chi^2_{2K}(\alpha)}{2U}$$

is itself a random variable, formed as the ratio of two random variables, since the $\chi^2_{2K}(\alpha)$ value can be considered a discrete random variable with the same number of realizable values as K, and, except for different sample values, the same probability law. (It is easily seen that for a given set of failure times, T_{ij}, the effect of permutation of the data is to alter the value of K, not U.) Thus, we start by considering the properties of the random variable λ_b when the system consists of two components.

A key step in Lieberman and Ross' paper is their defining a binomial random variable, J_i^*, where

$$J_i^* = \begin{cases} 1 & \text{with probability } \lambda_1/(\lambda_1+\lambda_2) \\ 2 & \text{with probability } \lambda_2/(\lambda_1+\lambda_2) \end{cases},$$

and showing that J_i^* has the same probability law as the expected failure pattern of the type 1 and 2 components. Thus, they show the unconditional probability density function for K is:

$$P\{K=k\} = \begin{cases} \binom{k-1}{k_1-1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k-k_1}, & k_1 \le k < k_2 - 1, \\ \binom{k-1}{k_1-1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k-1_1} & \\ \quad + \binom{k-1}{k_2-1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k-k_2}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k_2}, & k_2 \le k \le k_1 + k_2 - 1. \end{cases} \tag{7}$$

(Note there is an obvious misprint in this formula in the original paper.) The LR procedure arises from this observation, plus the well-known result that U has the conditional density function

$$f_{U|K=k}(u) = \frac{(\lambda_1+\lambda_2)^k}{(k-1)!}\, u^{k-1}\, e^{-(\lambda_1+\lambda_2)u}. \tag{8}$$

But, it follows immediately that λ_b has the conditional density function:

$$f_{\lambda_b|K=k}(\lambda_e) = f_{U|K=k}\left(\frac{\chi^2_{2K}(\alpha)}{2\lambda_e^2}\right)\cdot\frac{\chi^2_{2k}(\alpha)}{2\lambda_e^2}$$

$$= \frac{1}{\lambda_e}\,\frac{1}{(k-1)!}\left[\frac{\lambda\chi^2_{2k}(\alpha)}{2\lambda_e}\right]^k \exp\left\{-\lambda\chi^2_{2k}(\alpha)/2\lambda_e\right\},\ \lambda = \lambda_1 + \lambda_2, \tag{9}$$

and thus the conditional mean:

$$E(\lambda_b|K=k) = \int_0^\infty \lambda_e f_{\lambda_b|K=k}(\lambda_e)\,d\lambda_e$$

$$= \frac{1}{(k-1)!}\int_0^\infty \left[\frac{\lambda\chi^2_{2k}(\alpha)}{2\lambda_e}\right]^k \exp\left\{-\lambda\chi^2_{2k}(\alpha)/2\lambda_e\right\} d\lambda_e$$

$$= \frac{\lambda\chi^2_{2k}(\alpha)}{2(k-1)!}\int_0^\infty \tau^{k-2}\, e^{-\tau}\,d\tau,$$

where $\tau = \dfrac{\lambda\chi^2_{2k}(\alpha)}{2\lambda_e}$. Observe then:

$$E(\lambda_b|K=k) = \frac{\lambda\chi^2_{2k}(\alpha)}{2(k-1)} \quad , \quad k \geq 2 \qquad (10)$$

$$= \infty \qquad\qquad k = 1 \ .$$

Since the (unconditional) mean value of the upper bounds is:

$$E(\lambda_b) = \sum_{k=k_1}^{k_1+k_2-1} E(\lambda_b|K=k)P(k=k) \ , \qquad (11)$$

it follows that, unless $k_1 \geq 2$, $E(\lambda_b) = \infty$. In fact, the following proposition is an immediate consequence.

Proposition: For $k_1 \leq k_2$, the k_1^{th} moment of λ_b (the Lieberman and Ross upper bound to λ) will be unbounded.

Note also from (10) and (11), that when $k_1 \geq 2$,

$$E(\lambda_b) = \lambda\left\{\sum_{k=k_1}^{k_1+k_2-1} \frac{\chi^2_{2k}(\alpha)}{2(k-1)} P(K=k)\right\} = \lambda\, E([\chi^2_{2k}(\alpha)/2(k-1)]). \qquad (12)$$

It may be observed from the expression for $P(K=k)$ in equation (7) that $E(\lambda_b)$ depends, not on λ , but on λ_1/λ_2 , since $P(K=k)$ depends on this ratio. Thus we have shown:

Theorem: Given k_1 , k_2 and λ_1/λ_2 , the mean estimate for λ_b produced by the Lieberman and Ross technique will be the same constant percentage of λ , independent of the value of λ .

The r^{th} conditional moment of λ_b is given by

$$E(\lambda_b^r|K=k) = \lambda^r\left[\frac{\chi^2_{2k}(\alpha)}{2}\right]^r \frac{(k-r-1)!}{(k-1)!} \ , \ 1 \leq r \leq k \ . \qquad (13)$$

In particular,

$$\sigma^2_{\lambda_b|K=k} = \lambda^2 \left\{ \frac{\chi^2_2(\alpha)^2}{4(k-1)^2(k-2)} \right\}, \quad k > 2 . \tag{14}$$

Figure 1 shows $\{E(\lambda_b|K=k)/\lambda\}$ and $\{\sigma_{\lambda_b|K=k}/\lambda\}$ for a representative range of values of k. Note that for $k > 10$ the variation in the mean value is quite small. The above results easily generalize to the case of more than two component types.

As noted in the introduction, the "lost" data in the LR method is represented by those components which have not failed by U. (See equation (6)). Except in the case of simultaneous failure of all component types at U, there will be at least one component left of each type except one. Also, the LR method requires use of 100% of the data for one component type, hence the number of this type of component should not be included in any measure of relative data loss. Based on this discussion, we shall define the percentage data loss in any application of the LR method as

$$PL = \left\{ \sum_{i=1}^{N} (k_i - n_i) \right\} \Big/ \left\{ \sum_{\substack{i=1 \\ i \neq f}}^{N} k_i \right\}, \tag{13}$$

$$= NL \Big/ \left\{ \sum_{\substack{i=1 \\ i \neq f}}^{N} k_i \right\},$$

where f is the index of the component type which was 100% exhausted, i.e., $n_f = k_f$, and NL was defined by (6).

In the two component case, the mean of PL can be computed fairly straightforwardly from (7). Observe that the unconditional probability that the first type of component terminates the test is:

$$P(f=1) = \sum_{k=k_1}^{k_1+k_2-1} \binom{k-1}{k_1-1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k-k_1}$$

$$= \frac{1}{(k_1-1)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1} \sum_{k=0}^{k_2-1} \frac{(k_1+k-1)!}{k!}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k}. \quad (14)$$

But the conditional density function for n_2, given $f=1$, is:

$$P(n_2=k|f=1) = \binom{k+k_1-1}{k_1-1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k} [P(f=1)]^{-1}. \quad (15)$$

Thus the conditional expectation on n_2 is

$$E(n_2|f=1) = \sum_{k=0}^{k_2-1} k\binom{k+k_1-1}{k_1-1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1}\left(\frac{\lambda}{\lambda_1+\lambda_2}\right)^{k} [P(f=1)]^{-1}. \quad (16)$$

A similar expression is derivable for $E(n_1|f=2)$. Thus, based on (13), we see that the expected data loss for the two component case is:

$$E(PL) = \left\{\frac{k_1-E(n_1|f=2)}{k_1}\right\}P(f=2) + \left\{\frac{k_2-E(n_2|f=1)}{k_2}\right\} P(f=1)$$

$$= 1 - \frac{1}{k_1} E(n_1|f=2)P(f=2) - \frac{1}{k_2} E(n_2|f=1)P(f=1)$$

$$= 1 - \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k_2} \sum_{k=0}^{k_1-2} \binom{k+k_2}{k_2} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k+1}$$

$$- \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{k_1} \sum_{k=0}^{k_2-2} \binom{k+k_1}{k_1} \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k+1} \qquad (17)$$

(Note (17) is a function solely of k_1, k_2 and λ_1/λ_2 .)

Figure 2 shows curves for several combinations of k , and k_2 . Observe that each curve appears to have a unique minimum, occuring at or near the point

$$\frac{\lambda_2 k_1}{\lambda_1 k_2} = 1 \qquad (18)$$

i.e., when both subsystems have the same expected life before exhaustion. Therefore we also plotted the percentage of lost data as a function of

$$\frac{\lambda_2 k_1}{\lambda_1 k_2} \quad ,$$

and this is displayed in Figure 3. We would conjecture that the minimum data loss occurs when (18) is satisfied, and the curves indicate a monotonicity in terms of total number of components, but we cannot show this conjecture analytically.

Expanding (17) to cover $n > 2$, while possible, is computationally less useful since multiple series become involved.

III. DESCRIPTION OF THE SIMULATION FOR LR TEST

As part of this study, a simulation model was developed to generate individual component failure data, and perform the statistical analyses of the basis LR procedure. The programs were all written in FORTRAN IV, and run on the IBM 360/67 at NPS, using certain subroutines from both the NPS and IMSL libraries.

The simulation was developed to consider up to twenty different individual components of each of six different types. Individual components of the same type are assumed to be i.i.d, and the serial failure of all components of a type terminates the test. The simulation initially generates 120 shuffled random numbers, distributed exponentially with $\lambda=1$. The subroutine package LLRANDOM, as described in [2], is used. These numbers are then divided into blocks of twenty, and each block adjusted to produce exponentially distributed failures with $\lambda=\lambda_i$ $(i=1,2, \ldots, 6)$ by dividing by λ_i. Then, for each component type, cumulative failure data is constructed from the individual component failure times. From this data, the first component type to be exhausted is determined, and the number of components of each of the other types that have failed up to that time, is also determined. With this data, the quantities U and $2K$ needed for the LR test are determined, and the estimated bound for λ is computed using the subroutine PRCHI from the IMSL package to compute the inverse chi-square table.

At this time, the routine also collects statistics on the unused data. Specifically, it computes the percentage of usable data lost, defined by (13). Repeated iterations (normally 300) of the simulation are run for each choise of k_i and λ_i. At the completion of these iterations,

additional statistics are compiled on the mean, variance and 100α% level of the estimates on reliability obtained from the individual trials.

Initial testing and validation of the simulation was carried out for the two component (N=2) case. Runs were made with various combinations of values for λ_1, λ_2, k_1 and k_2 (Table 1). Each run consisted of 300 iterations of the LR procedure, where each iteration consisted of generating one set of T_{ij}'s and the corresponding LR estimate for λ. Each set of pseudo-random failure times was used only once, and without any reordering. Validation consisted of comparing the results of several simulations to predicted values. First was the comparison of the 100(1-α)% estimate of reliability to the true reliability. Since the LR procedure is an exact bound for λ, and since, in these tests, we used an $\alpha=.95$ confidence level, we expected that on each run, 95% of the bounds would fall above the true lambda. Thus, after each iteration, we ordered the bounds on λ in increasing order, and chose the 15th (out of 300). The ratio of this value to the true reliability, $\lambda_t = (\lambda_1 + \lambda_2)$ was computed, and compared to unity. Figure 4 shows, as a function of λ_t, the scatter diagram of ratios obtained for runs.

The second comparison made was between the mean value of the λ_b that were computed in the simulation and the theoretical mean predicted by equation (12). The results of this comparison are shown in Figure 5, where the λ_t (the actual λ used to generate the data) is plotted on the horizontal axis, and the vertical axis displays the ratio:

$$\bar{\lambda}_b / E(\lambda_b)$$

where: $\bar{\lambda}_b$ is the (arithmetic) mean of the computed upper bounds and

$E(\lambda_b)$ is the theoretical mean, (12).

The final comparison made was between the average actual data loss in a run, and the expected data loss predicted by (17). The result of this comparison are shown in Figure 6, where λ_t is on the horizontal axis and $(\overline{PL}/E(PL))$ is on the vertical. Here $\overline{PL}$ denotes the (arithmetic) mean percentage of data lost in the iterations of a run, and $E(PL)$ is as given in (17).

As can be seen from Figures 4-6, the simulated values were all in acceptable agreement with the predicted ones, and therefore, we concluded the simulation itself is valid.

IV. THE EFFECT OF LOST DATA ON THE LR METHOD

With the simulation, as described in Section III debugged and validated, we proceeded to investigate the effect of lost data on the LR method. This investigation proceeded along two major lines. First was consideration of the effect when no prior ordering of the data was used, i.e., the basic LR method. In this instance we were particularly interested in any relationship that might exist between the percentage of data lost and the accuracy of the bound, λ_b , derived. Our second consideration was the effect on the estimated bounds of a priori ordering of the data, e.g., ordering in order of increasing inter-failure times. As noted in the introduction, such methods will produce biased estimates for the bounds, and a major concern is whether the variance of the bounds is smaller than that of the standard LR method, and whether the bias is predictable.

To investigate the relation between the lost data and the accuracy of the LR bound, the simulation was programmed to produce, at the end of each run, a scatter diagram showing the relative accuracy of the estimate (λ_b/λ_t) versus the percentage of lost data on each iterations. The number of iterations was increased so that each run consisted of 500 iterations. Several representative such diagrams are shown at Figures 7-12. Although each case evidenced some "tightening" of the group of estimates as the percentage of lost data decreased, the magnitude of this effect is really noticeable only when the total sample is small. This, of course, relates to the fact that the most rapid changes in Figure 1 occur for $K \leq 7$. It is also worth noting that the amount of scatter tended to decrease for $(k_1\lambda_2/k_2\lambda_1) \simeq 1.0$. Again this seems intuitively clear.

To study the effect of ordering, the program was modified so that after the individual component interfailure times were generated, they were rearranged into increasing order. This clearly will produce a biased high upper bound, since it, in effect, includes only the least reliable components in satisfying:

$$\sum_{k=1}^{n_i} T_{ij} \leq U \ .$$

Figures 13-15 display the consolidated results of 370 runs of 300 iterations each, with a run corresponding to different values of k_1, k_2, λ_1 and λ_2. Note that in each of these graphs the horizontal variable is chosen as

$$L_{12} = \min\left\{\frac{\lambda_1 k_2}{\lambda_2 k_1} \ , \ \frac{\lambda_2 k_1}{\lambda_1 k_2}\right\} , \tag{19}$$

rather than $\lambda_t = (\lambda_1 + \lambda_2)$. The reason for this was that, after some reflection, we felt that the reordering described here should have decreasing effect as L_{12} deviated from unity. (Note that L_{12} simply reflects the ratio of expected failure times for all the components of each individual type.) The general trend in these figures is consistent with our expectations, i.e., the bounds for λ (denoted λ_b^o) are consistently biased higher than their counterparts in the unordered case (Figures 13-14), and the average data loss is consistently less than in the unordered case. However, on close inspection, observe that the dispersion in Figure 13 seems more marked than that in the unordered case (Figure 4), and more pronounced near L_{12}=0. This we had not expected, and therefore was investigated in some more detail.

After some consideration, we decided that the accentuated spreading (especially near $L_{12}=0$) arose primarily from the tendency of the exponential distribution to produce some highly unreliable components (i.e., $T_{ij} \sim 0$), which could produce a marked variation in estimates for tests with small numbers of components. To confirm that this apparent increased spreading is actual, in Figures 16-17 we compared dispersions of the estimates in the ordered scheme with those in the unordered scheme. Specifically, for each iteration of each run, bounds for λ_t were produced, both using prior ordering (denoted by λ^o_{bi}), and without prior ordering (denoted λ^u_{bi}). Then, for each run the ratios of actual magnitudes of variances observed:

$$R_m = \frac{\left\{E([\lambda^o_{bi}]^2) - E^2(\lambda^o_{bi})\right\}}{\left\{E([\lambda^u_{bi}]^2) - E^2(\lambda^u_{bi})\right\}} , \tag{20}$$

and the ratio of relative variances:

$$R_r = R_m\left[\frac{\bar{\lambda}^u_b}{\bar{\lambda}^o_b}\right]^2 \tag{21}$$

were computed, where the usual estimates are used, e.g.,

$$E(\lambda^u_{bi}) = \bar{\lambda}^u_{bi} = \frac{1}{300}\sum_{i=1}^{300} \lambda^u_{bi} . \tag{22}$$

Finally, as a function of L_{12}, we plotted $\sqrt{R_m}$ (Figure 16) and $\sqrt{R_r}$ (Figure 17). As expected, for all intents, $\sqrt{R_m}$ always exceed unity. That is, the actual variance of the ordered scheme exceeded that of the

unordered scheme. Furthermore, the relative variance of the ordered scheme exceeded even the relative variance of the unordered scheme a significant portion of the time when $L_{12} > 0.5$. As we noted, for small L_{12} the expected effect of ordering should be less noticeable, and we attribute the fact that the relative variance here, as given by (21), did not exceed unity primarily to the fact that $\bar{\lambda}_b^o > \bar{\lambda}_b^u$.

Figures 13-17 do not seem to indicate that a priori arrangement of the component failure data in order increasing interfailure time offers any significant improvement over random ordering, largely because the variances displayed in Figure 14 suggest that prediction of the bias introduced by the ordering is not predictable. (Actually, this comment must be strongly qualified, since in Figure 14 we have expressed the variance based on the single parameter, L_{12} . It is still quite possible that a more predictable relation could emerge were we to retain k_1, k_2 and (λ_1/λ_2) as three independent parameters.)

We also feel that any other a priori ordering algorithm will be no more successful in producing lower variance bounds as long as one attempts to predict the bias using a single parameter. There are, however, two other possibilities that should be investigated. One, as noted above, is to include several parameters in the prediction. The second is to recognize that every different ordering of the data produces an estimated bound, and each is, statistically, equally valid. This observation means that each LR test produces data which yield several samples from the same population. The implication is then that one should consider all the possible bounds that can be estimated from all the different possible orderings of the data. Thus, for example, the data in section I should be

viewed as yielding that both 6.30 and 4.75 are 95% confidence (upper) bounds on λ, and this is certainly stronger than the statement that either one alone is such a bound. (Note it might be argued that this will produce an unmanageably large number of samples when the number of failed components is large. However, as reference to Figure 1 indicates, where the number of failed components is large, the variance on all the bounds is small, and so some randomly chosen smaller sample could be used.)

V. SUMMARY AND CONCLUSIONS

Given a set of component interfailure times to be converted into an estimate on system reliability by the Lieberman-Ross technique, it is well-known that one parameter ("U") needed for the procedure is uniquely determined, irrespective of the order of the data, while the value of the second parameter ("K") is data order dependent and not uniquely determined. Furthermore, the design of the LR procedure requires that, in general, some relevant, valid data be discarded. The purpose of this paper was to consider the effects of this "lost" data and the data order.

Since the data order determines the degrees of freedom for the chi-square test, we considered the distribution of estimated bounds on λ produced for a fixed degree of freedom. Formulas were presented for the mean and variance of these bounds, and these were shown to be constant percentages of λ for fixed K. We also programmed a simulation to investigate the impact of a priori ordering of the failure data. The ordering we chose was in order of increasing interfailure time (least reliable first), which is equivalent to including maximum data. The results of this simulation showed that this a priori ordering not only produced biased bounds for λ, but, expressed in terms of the single parameter $(\lambda_1 k_2 / \lambda_2 k_1)$, these estimated bounds were more dispersed than the estimates produced without prior ordering. Thus we concluded it appears that a priori ordering offers no improvement over the basis LR procedure, as long as one attempts to predict the resultant bias in terms of a single parameter. In future investigations we hope to consider both multi-parameter prediction of the prior ordering bias, and improvement in the estimated bounds possible by utilizing several reorderings of the data.

ACKNOWLEDGEMENTS

This research was supported by the Foundation Research Program of the Naval Postgraduate School, which is funded by the Chief of Naval Research. The author is indebted to Professor Toke Jayachandran, for originally suggesting this investigation, and for several very helpful comments during it.

REFERENCES

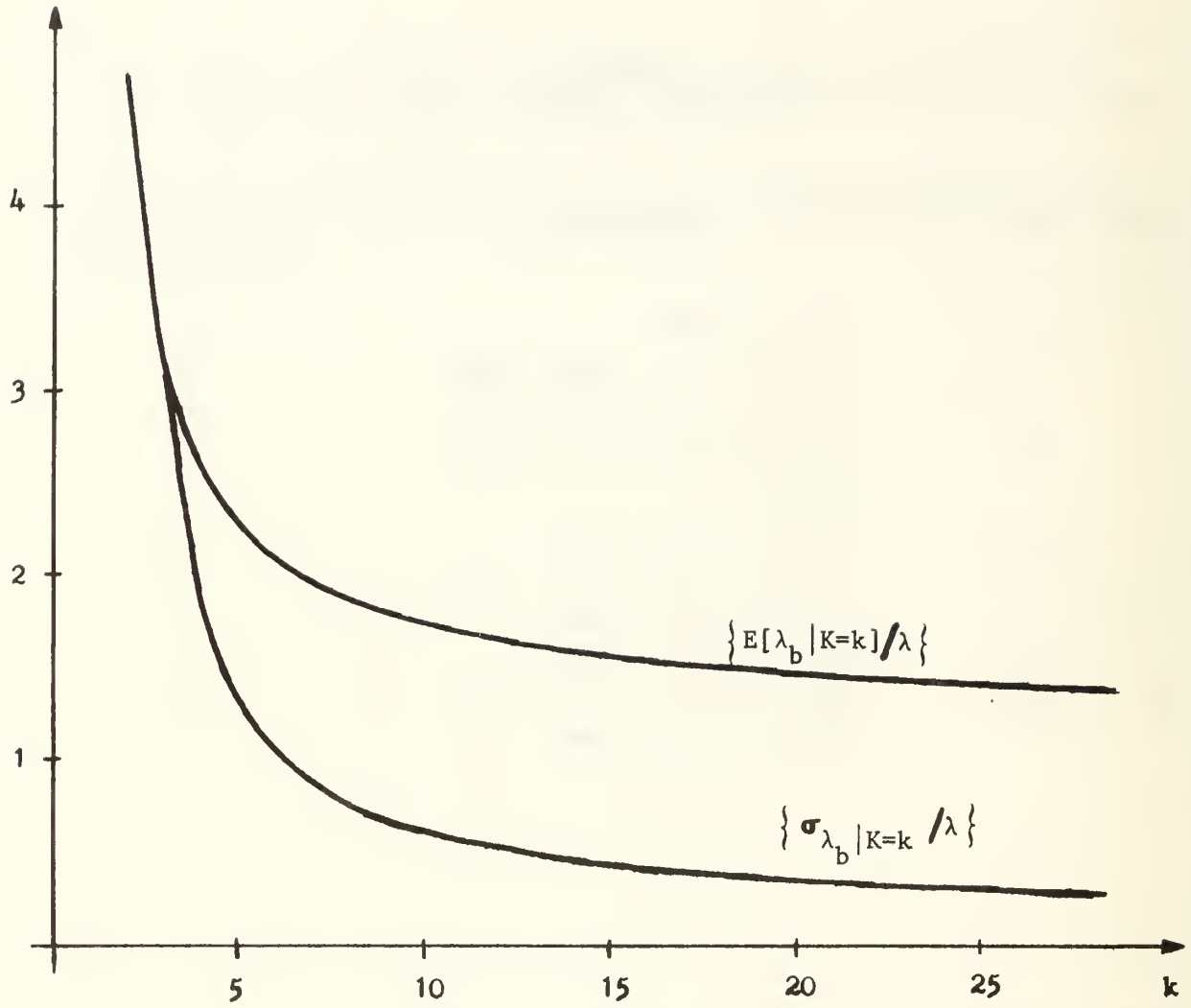
[1] Lieberman, J. and Ross, S., "Confidence Intervals for Independent Exponential Series Systems," J. Am. Statistical Assn., Vol. 66, Number 336, December 1971, pp 837-840.

[2] Learmonth, G. and Lewis, P.A.W., "Naval Postgraduate School Random Number Generator Package LLRANDOM," NPS Report 55LW73061A, June 1973.

TABLE 1

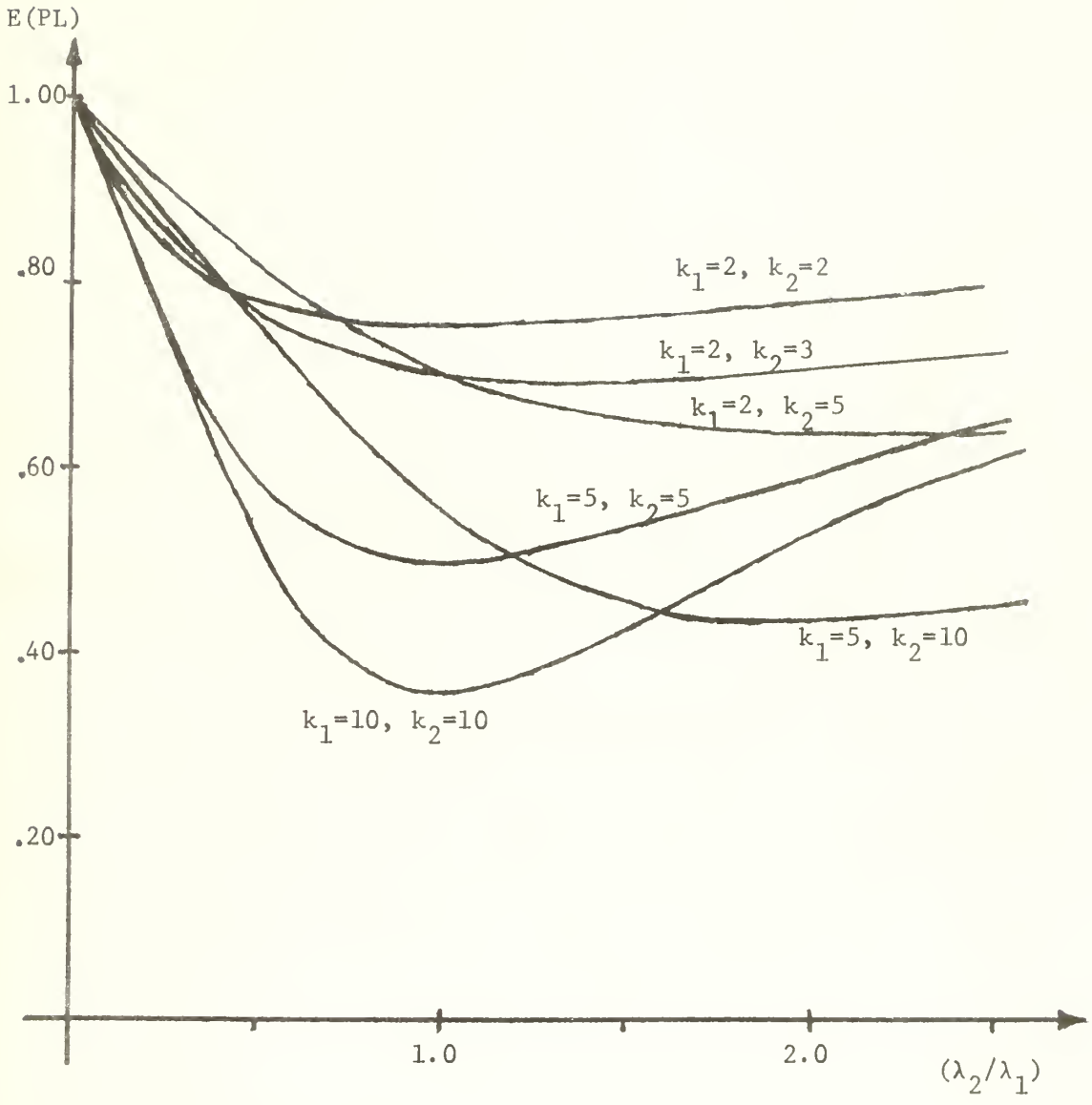
k_1	k_2	VALUES OF λ_t	RUNS AT EACH λ_t*
2	2	0.001, 0.005	20
2	3	0.001, 0.005, 0.010, 0.020	10
2	5	0.001, 0.005, 0.010, 0.020	10
3	3	0.001, 0.005, 0.010, 0.020, 0.050	10
3	5	0.020, 0.050, 0.100	10
5	5	0.001, 0.005, 0.010, 0.020, 0.050, 0.100, 0.200	10
5	10	0.050, 0.100, 0.200	10
10	10	0.001, 0.005, 0.010, 0.020, 0.050, 0.100, 0.200	10

*Each run consisted of 300 iterations of an LR failure test using fixed, but randomly chosen values of λ_1, λ_2, subject to $\lambda_1 + \lambda_2 = \lambda_t$.



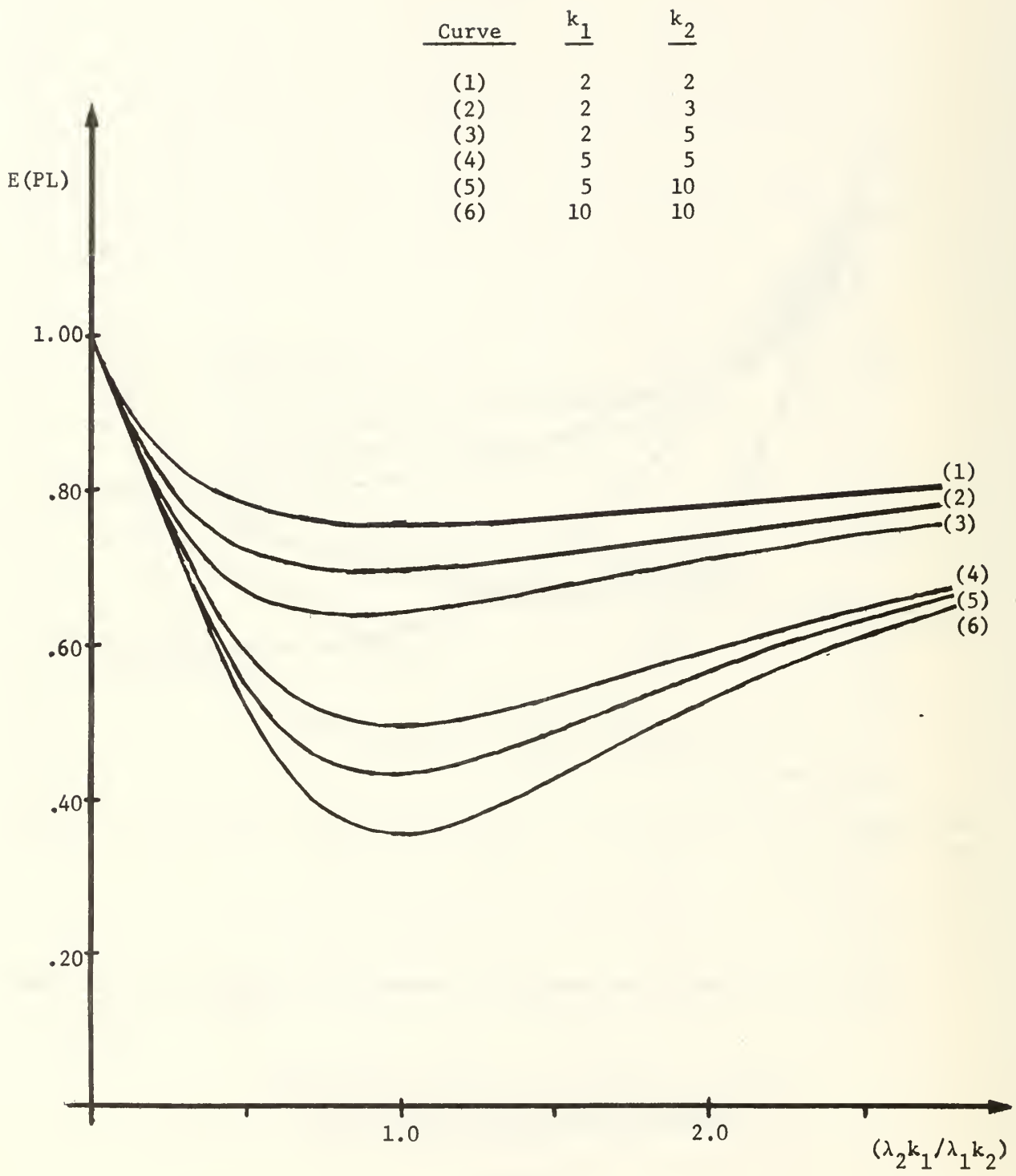
Dependence of Mean and Variance of Bound
on Data Size

Figure 1



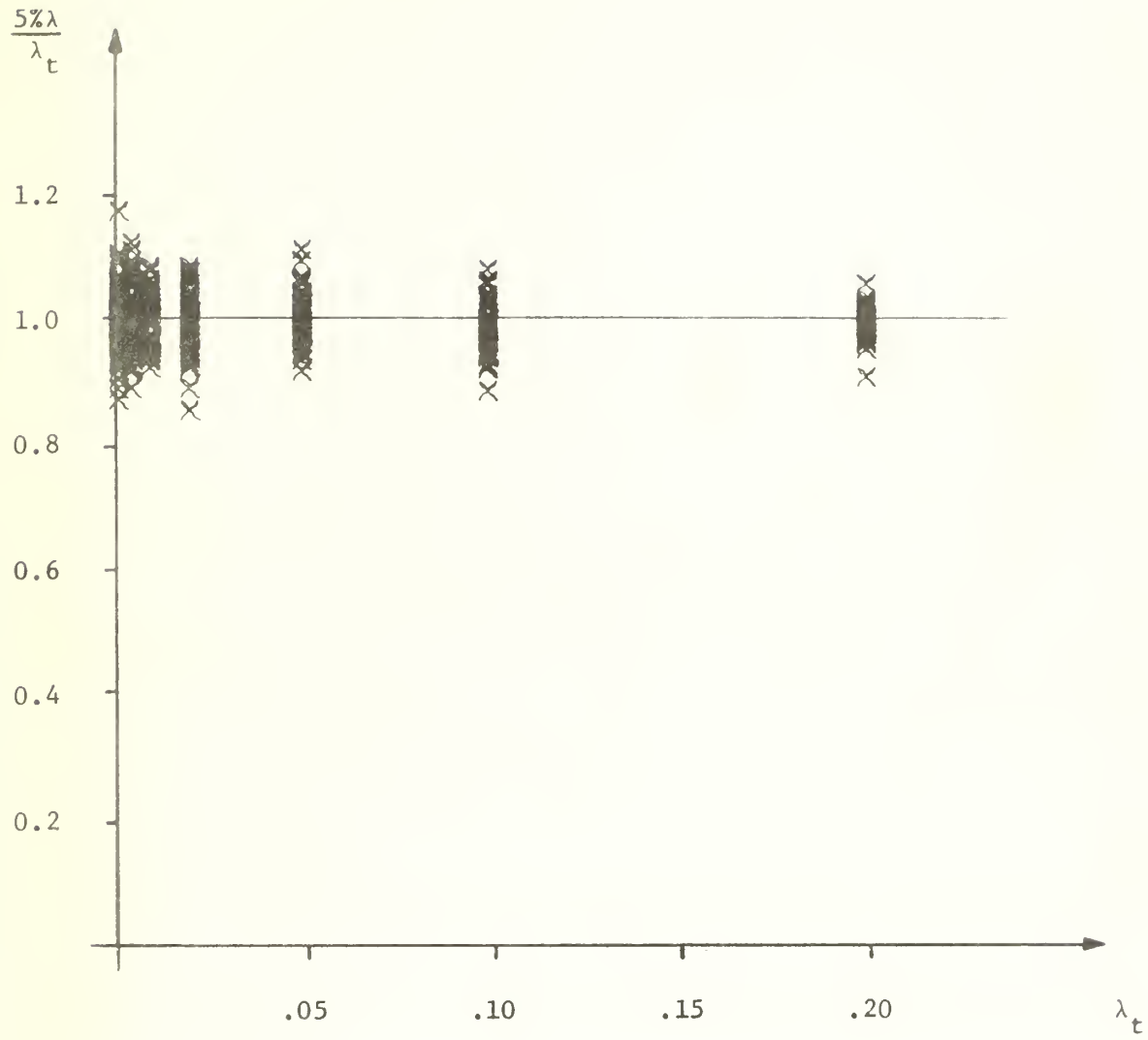
Expected Percentage of Data Lost

Figure 2



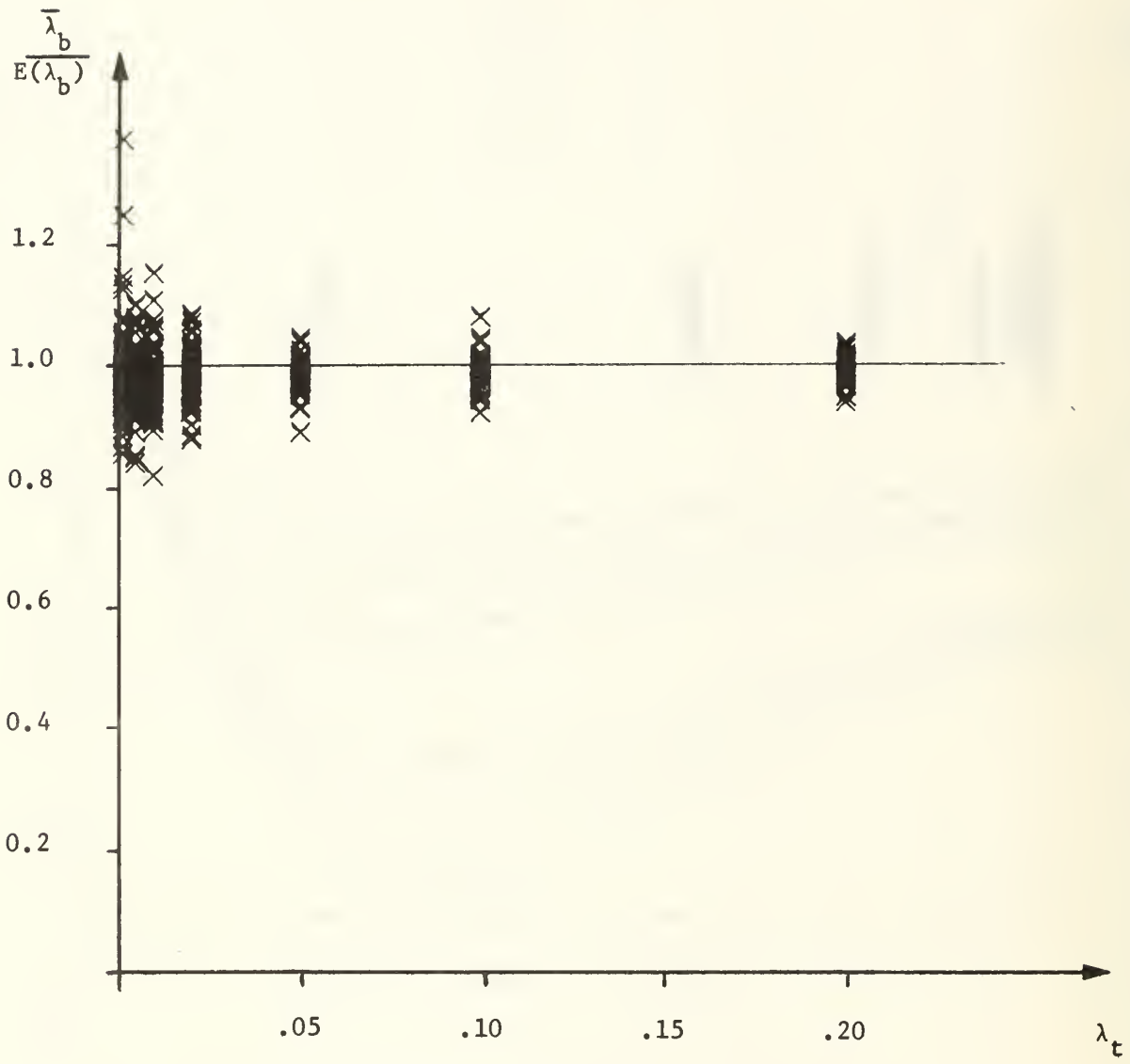
Expected Percentage of Data Lost

Figure 3



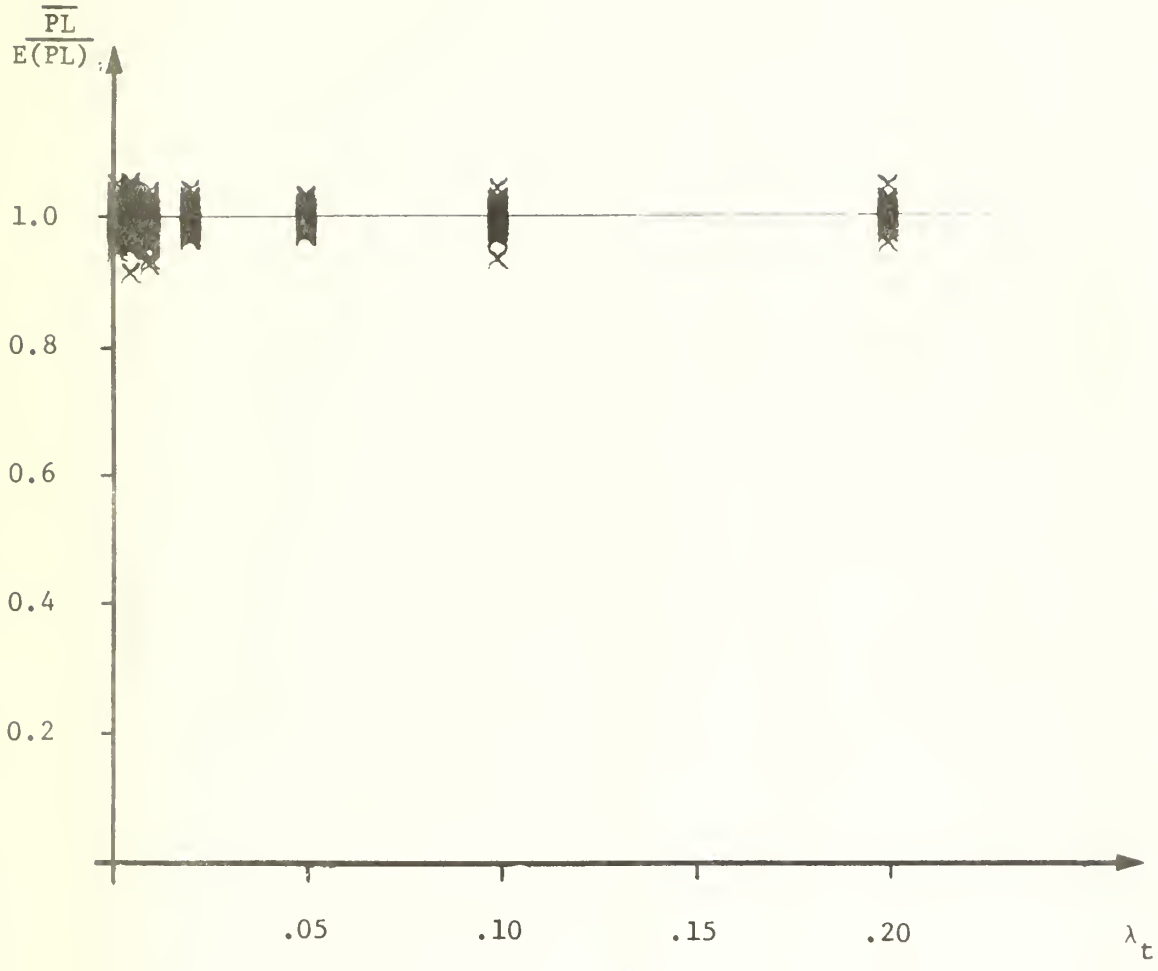
Test of Simulation For Statistical Exactness

Figure 4



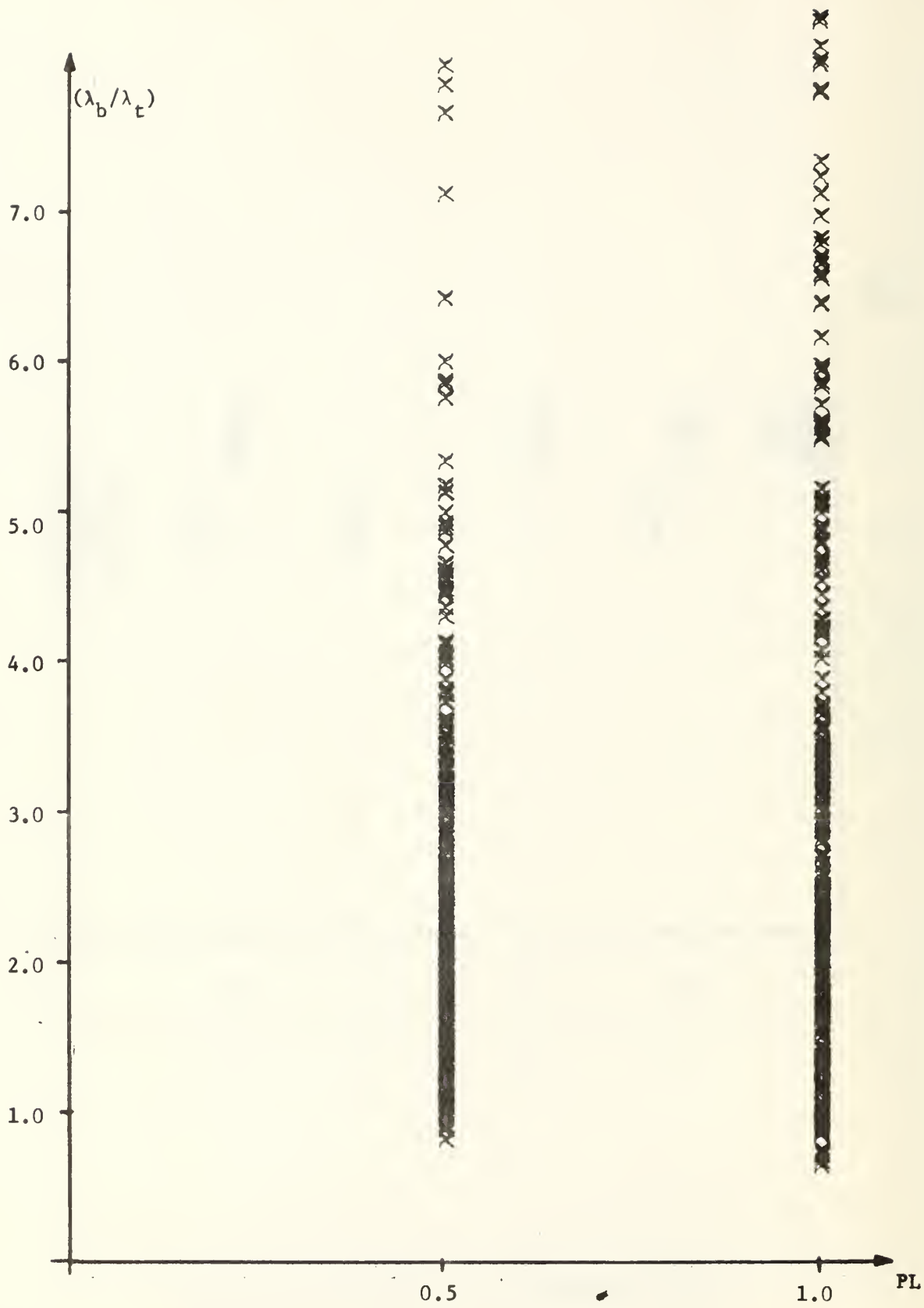
Test of Simulation for Predicted Mean Bound

Figure 5



Test of Simulation for Predicted Lost Data

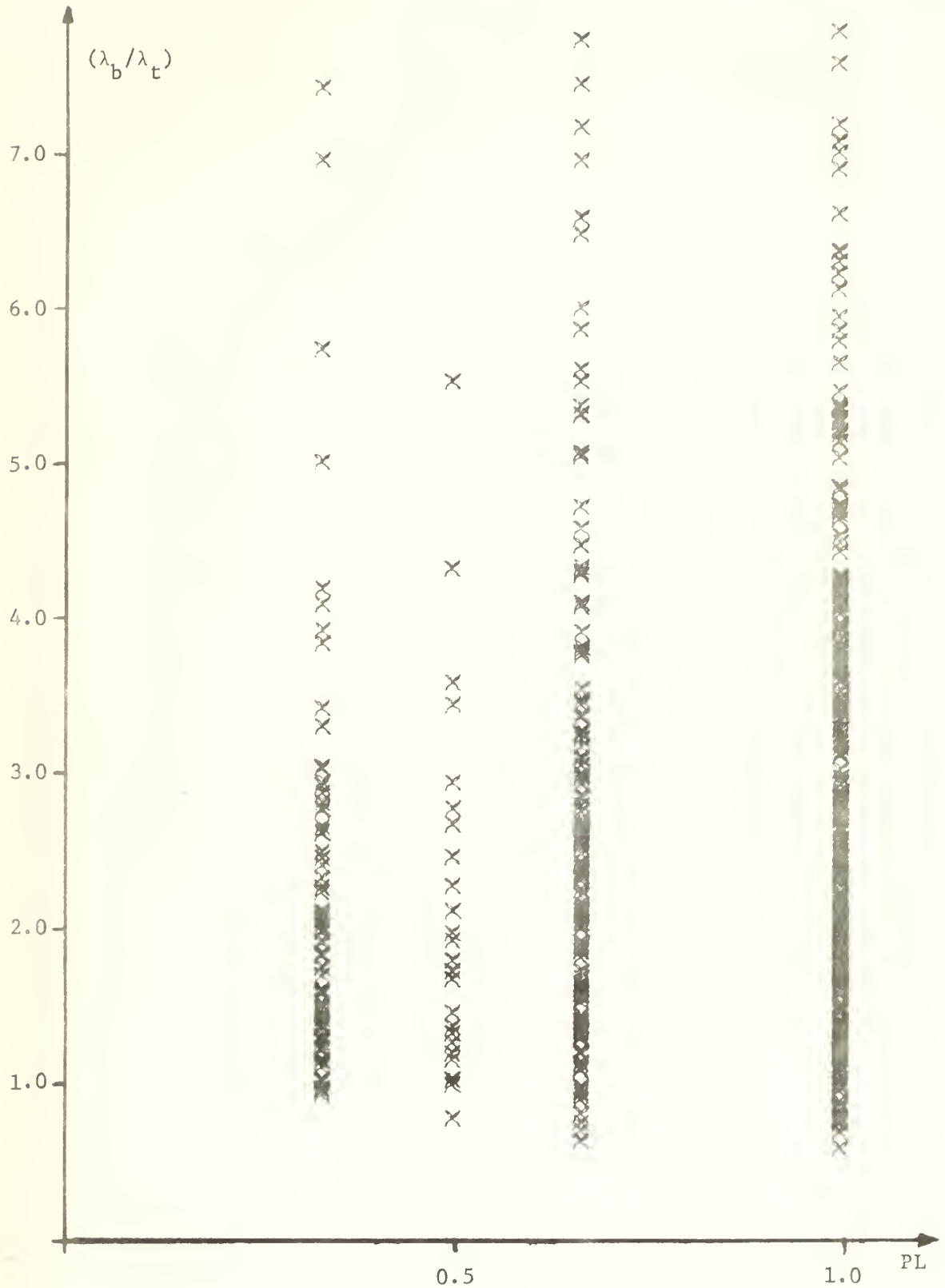
Figure 6



Computed Bound (λ_b) Versus Percentage of Data Lost (PL)

$k_1 = 2$, $\lambda_1 = .0005$, $k_2 = 2$, $\lambda_2 = .0005$

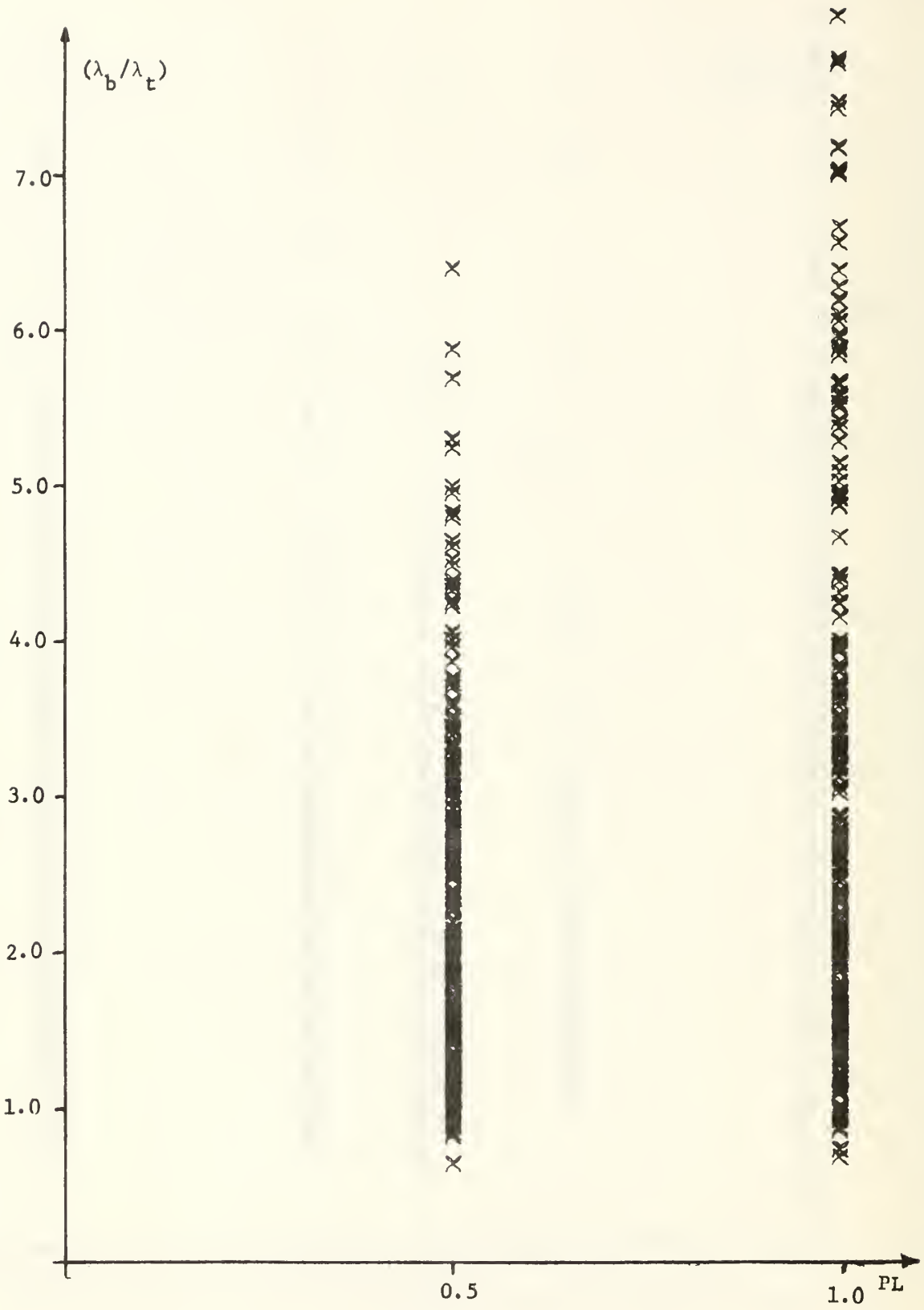
Figure 7



Computed Bound (λ_b) Versus Percentage of Data Lost (PL)

$k_1 = 2$, $\lambda_1 = .0007$, $k_2 = 3$, $\lambda_2 = .0003$

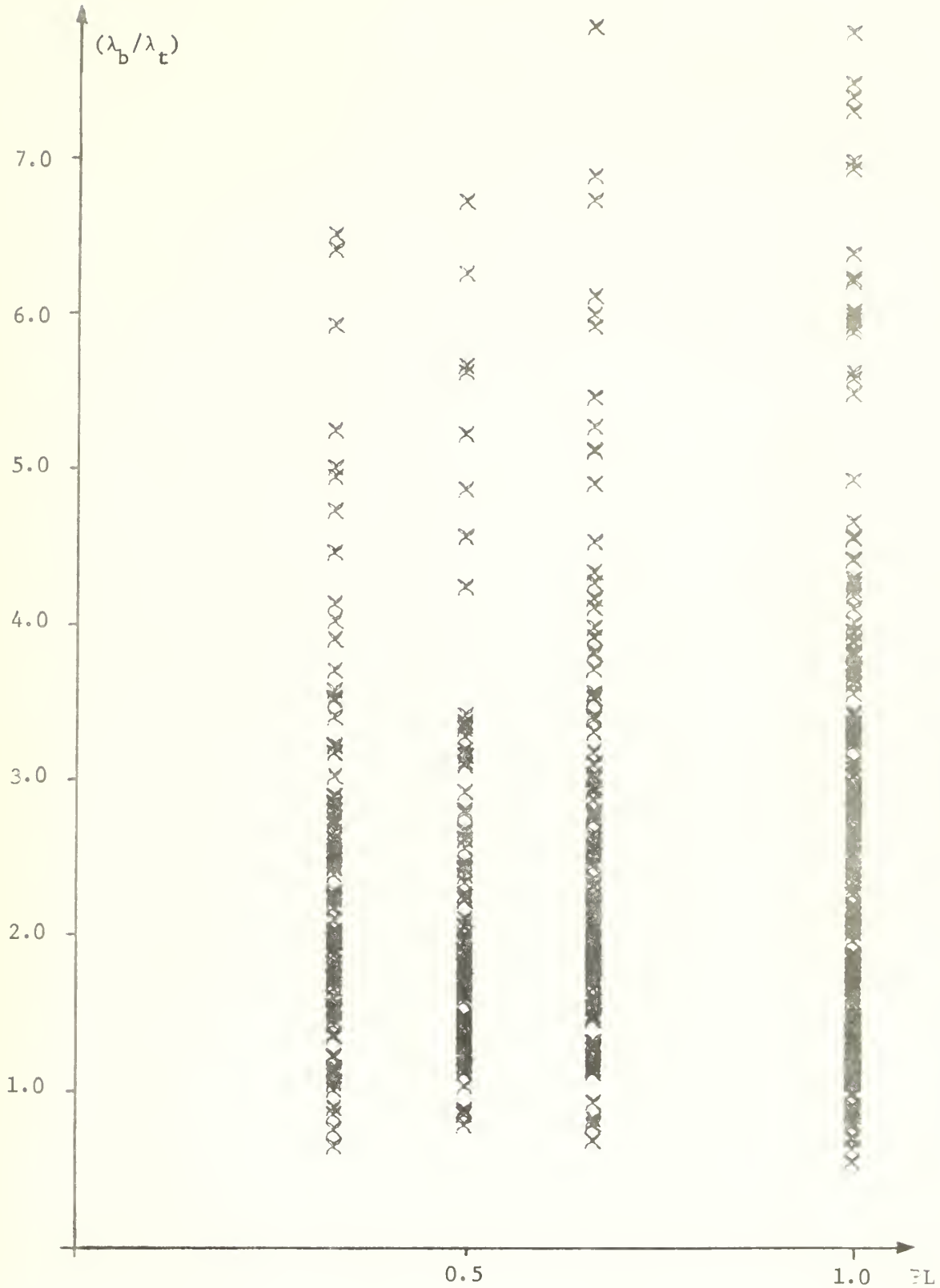
Figure 8



Computed Bound (λ_b) Versus Percentage of Data Lost (FL)

$k_1 = 2$, $\lambda_1 = .0007$, $k_2 = 2$, $\lambda_2 = .0003$

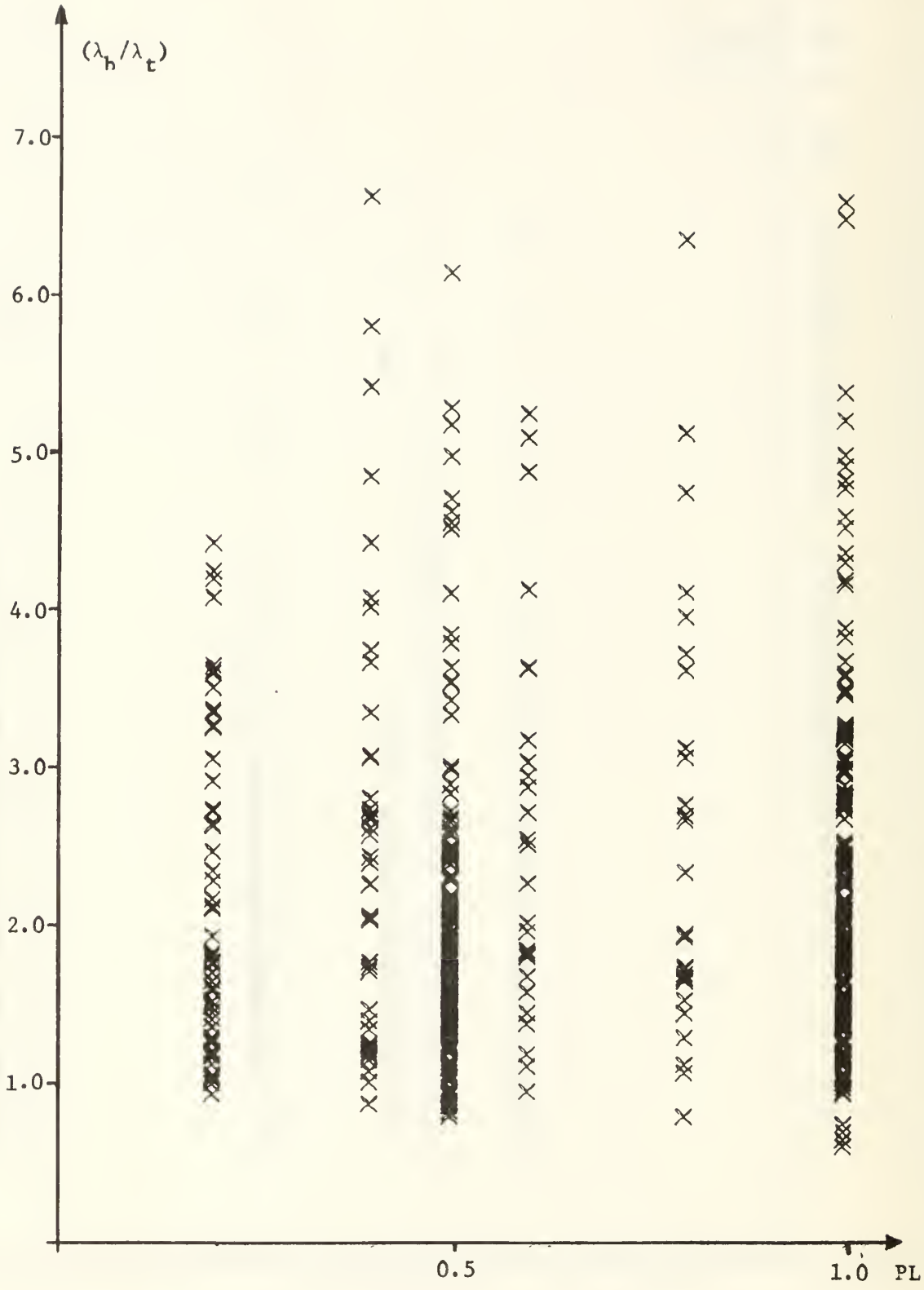
Figure 9



Computed Bound (λ_b) Versus Percentage of Data Lost (PL)

$k_1 = 2$, $\lambda_1 = .0005$, $k_2 = 3$, $\lambda_2 = .0005$

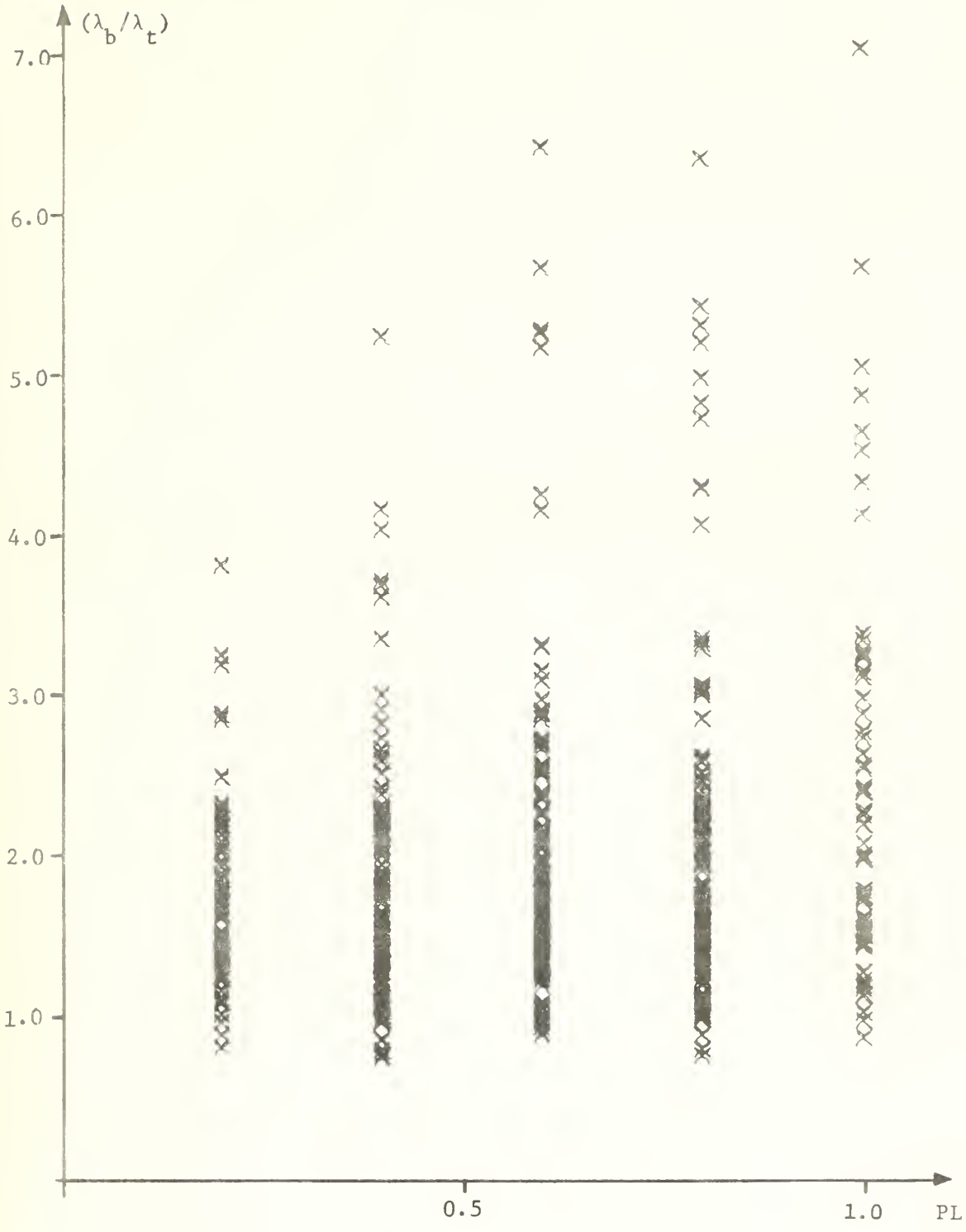
Figure 10



Computed Bound (λ_b) Versus Percentage of Data Lost (PL)

$k_1 = 2$, $\lambda_1 = .0002$, $k_2 = 5$, $\lambda_2 = .0008$

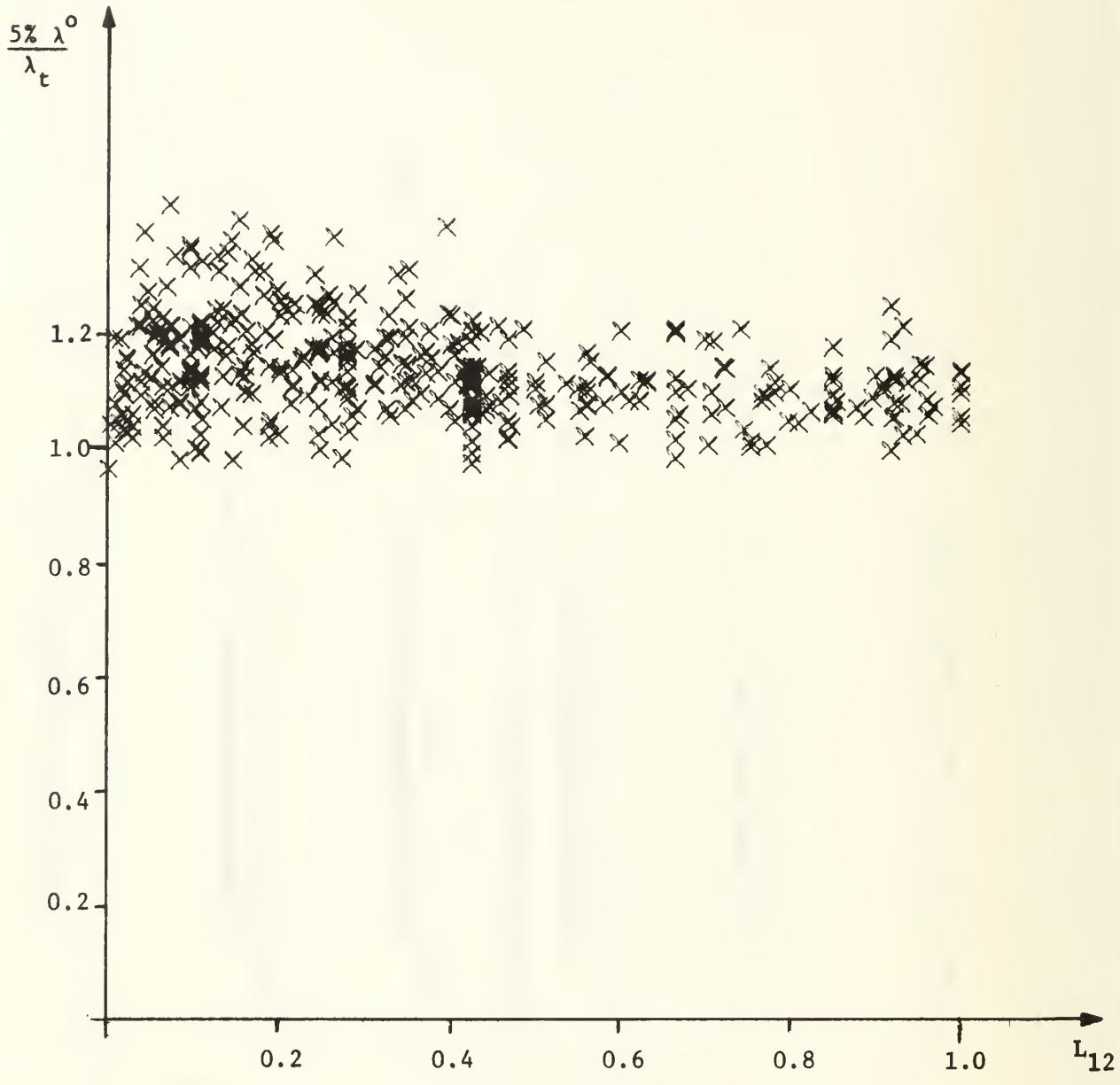
Figure 11



Computed Bound (λ_b) Versus Percentage of Data Lost (PL)

$k_1 = 5$, $\lambda_1 = .0007$, $k_2 = 5$, $\lambda_2 = .0003$

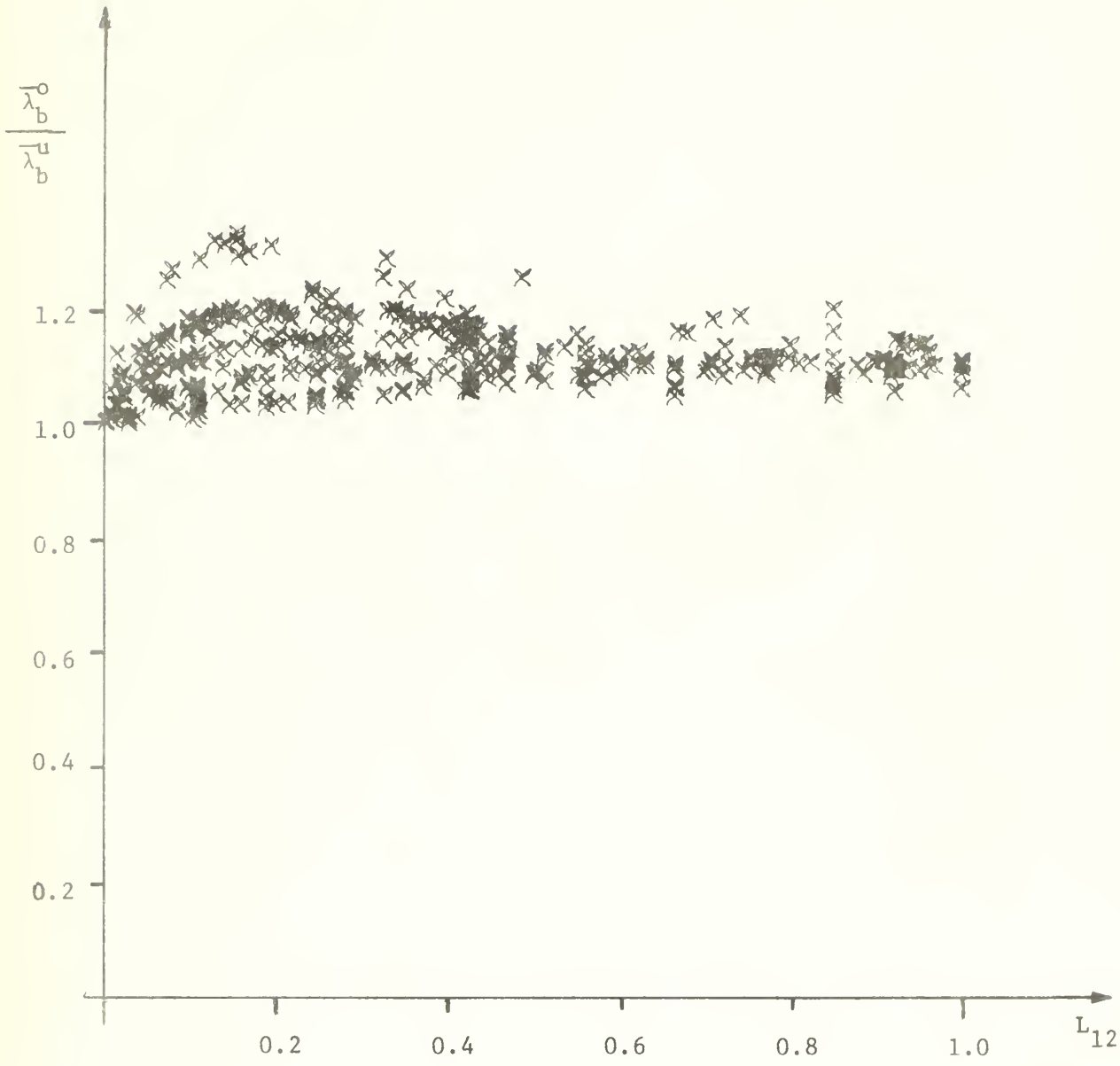
Figure 12



Ratio of 5% Estimated Bound to True λ

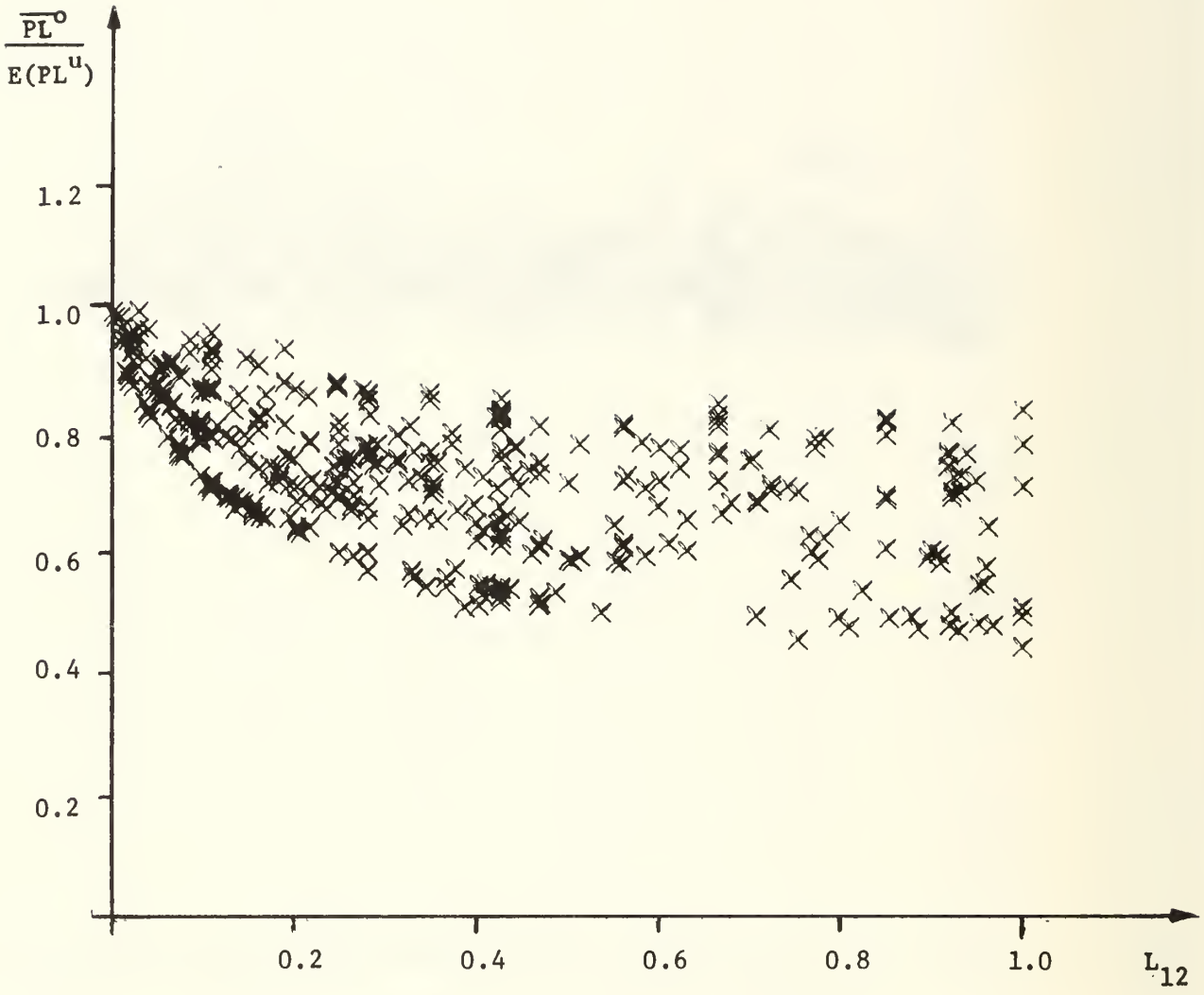
Prior Ordering

Figure 13



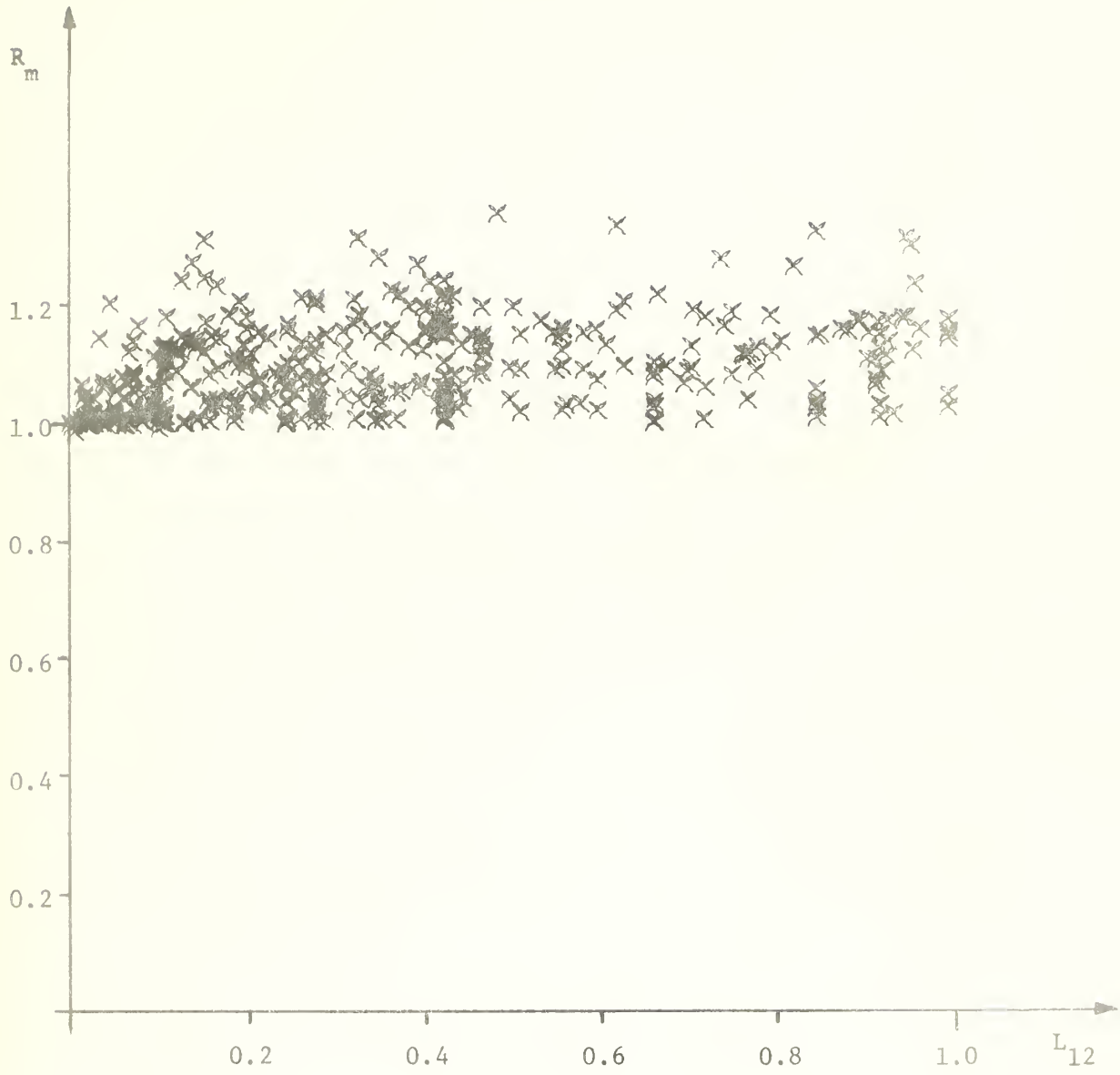
Ratio of Mean Estimated Bounds with $(\overline{\lambda_b^o})$ and without $(\overline{\lambda_b^u})$ Prior Ordering

Figure 14



Ratio of Mean Percentage of Data Lost with
Prior Ordering ($\overline{PL^o}$) to Expected Loss without Ordering ($E(PL^u)$)

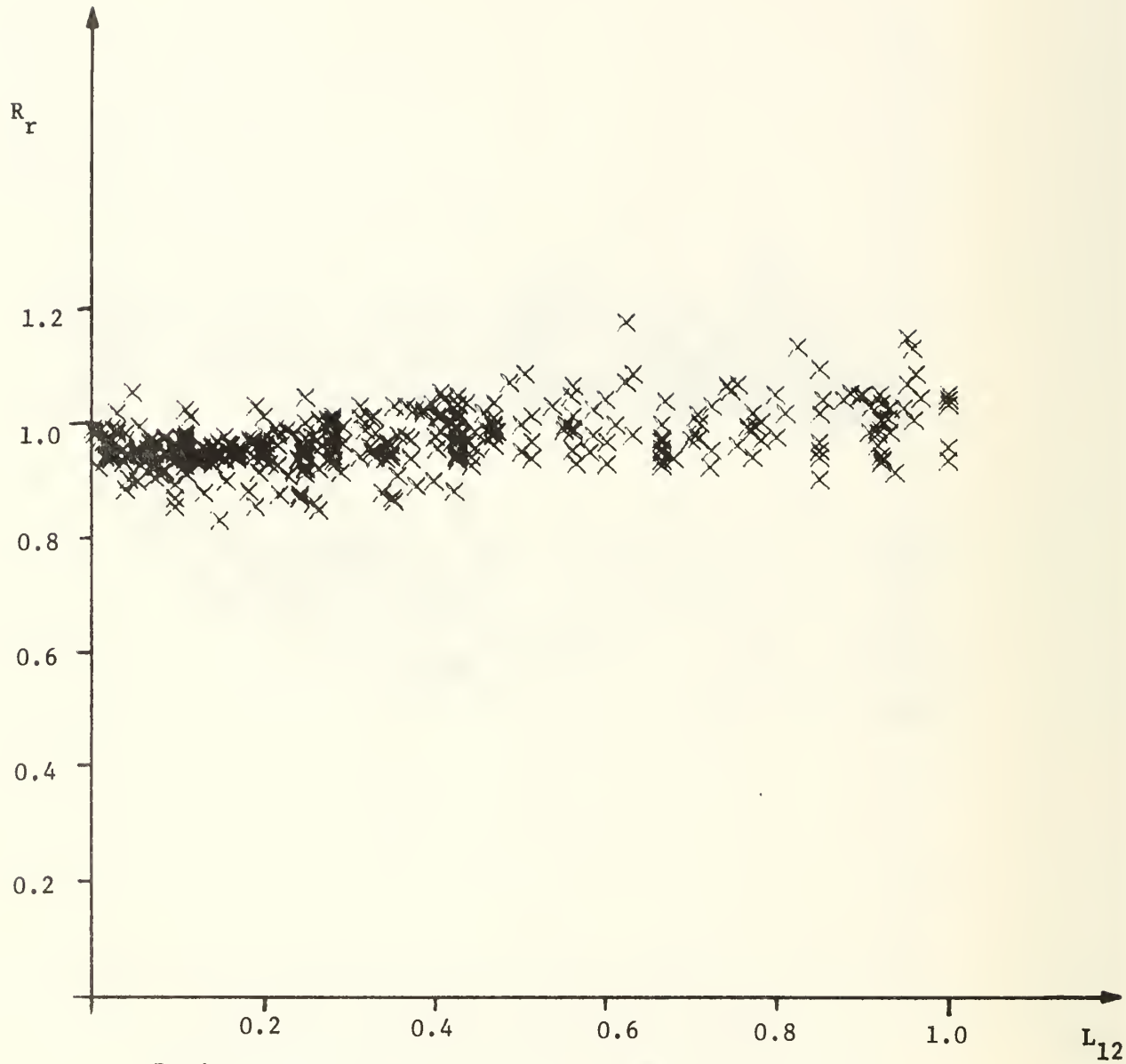
Figure 15



Ratio of Actual Magnitudes of Standard Deviations

With and Without Prior Ordering

Figure 16



Ratio of Relative Magnitudes of Standard Deviations
With and Without Prior Ordering

Figure 17

DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	12
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Dr. A.L. Schoenstadt, Code 53Zh Department of Mathematics Naval Postgraduate School Monterey, California 93940	10
4.	Chief of Naval Research Department of the Navy 800 N. Quincy St. Arlington, Virginia 22217	2
5.	Dean of Research Naval Postgraduate School Monterey, California 93940	2
6.	Professor Toke Jayachandran Office of Naval Research, Code 431 800 N. Quincy St. Arlington, Virginia 22217	3
7.	Professor J. Esary, Code 55Ey Naval Postgraduate School Monterey, California 93940	1
8.	Dean W.M. Woods, Code 500 Naval Postgraduate School Monterey, California 93940	1
9.	Dr. Janet Myhre Mathematics Department Claremont Men's College Claremont, California 91711	1